

# Homework 3

## Introduction

In this homework it is the idea to use the simulated data we introduced before, in conjunction with the R “SNPassoc” package, to answer relevant questions in a population-based genetic association context. However, if you can answer the questions using another technique you have seen in other classes, this is fine as well.

## Specific questions on population-based genetic association analysis

Q1. Which individuals in the data should you use for the population-based genetic association analyses we envisage?

Q2. Do a genome-wide check on the validity of the HWE condition. Which markers do not satisfy this condition? [Note: Use a function which is able to check this condition using one line only (it is not the idea that you manually perform this for each marker separately, as in a previous homework)]

Q3. Are there genotypes missing? Do you observe particular patterns in the observed/missing genotypes?

Q4. Are there monomorphic SNPs? How should they be treated further on in subsequent analyses?

Q5. In the phenotype-file, a continuous trait has been provided. Consider the following two scenarios:

- Scenario 1: take the provided continuous trait and analyze it as such for its association with the provided markers, using population-based genetic association analysis techniques to identify important genetic main effects
- Scenario 2: dichotomize the continuous trait by computing the upper extreme 20% cut-off (threshold) trait value based on parental phenotypes. Values above this threshold are recoded as 1, values below this threshold are recoded as 0. Use the new binary trait in a population-based genetic association analysis, using all markers as potential main effects.

Compare the results from both scenarios (team work – for final project). What do you observe? Is this in line with expectations?

Q6. Compute appropriate significance levels for your association tests in Q5. Try out several methods and compare / discuss your results.

Q7. Perform a subset analysis for males / females on the quantitative trait. Do you see any differences? Was this to be expected?

Q8. Perform a gene-gene interaction analysis for the quantitative trait, using the tools provided by the “SNPassoc” package. Do you observe any interesting signals?

Write a small report, including some explanations about how you obtained the answers

**Due date:** 1 April 2011